



Gp41-engineered loop = SEQ ID NO 7 and NO 8

M Q A R Q L L S G I V Q Q Q N N L L 18
ATG CAG GCC AGA CAA TTA TTG TCT GGT ATA GTG CAG CAG AAC AAT TTC CTG
R A I E A Q Q H L L Q L T V W G I K 36
AGG GCT ATT GAG GCG CAA CAG CAT CTG TTG CAA CTC ACA GTC TGG GGC ATC AAG
Q L Q A R I L A V E R Y L K D Q Q L 54
CAG CTC CAG GCA AGA ATC CTG GCT GTG GAA AGA TAC CTA AAG GAT CAA CAG CTC
L G I D G S S G G R G G S N A S W S 72
CTG GGG ATT GAC GGT AGC AGT GGA GGT AGA GGT GGA TCC AAT GCT AGT TGG AGT
N K S L E Q I W N H T T W M E W D R 90
AAT AAA TCT CTG GAA CAG ATT TGG AAT CAC ACG ACC TGG ATG GAG TGG GAC AGA
E I N N Y T S L I H S L I E E S Q N 108
GAA ATT AAC AAT TAC ACA AGC TTA ATA CAC TCC TTA ATT GAA GAA TCG CAA AAC
Q Q E K N E Q E L L E L D L E H H 126
CAG CAA GAA AAG AAT GAA CAA GAA TTA TTG GAA TTA GAT CTC GAG CAC CAC CAC
H H H * 129
CAC CAC CAC TGA

FIGURE 1

2A

Gp41 (region 545-684): SEQ ID NO 1

QARQLLSGIVQQQNNLLRAIEAQQHLLQLTVWGIK
QLQARILAVERYLKDQQLLGIWGC SGKLICTTAVP
WNASWSNKSLEQIWNNMTWMEWDREINNYTSLIHS
LIEESQNQQEKNEQEELLELDKWASLWNWFNTNWL

2B

Gp41 regions

Region 555-577	SEQ ID NO 3	QQQNNLLRAIEAQQHLLQLTVWG
Region 572-601	SEQ ID NO 4	QLTVWGIKQLQARILAVERYLKDQQLLGIW
Region 590-620	SEQ ID NO 5	RYLKDQQLLGIWGC SGKLICTTAVPWNASWS
Region 628-663	SEQ ID NO 6	WNNMTWMEWDREINNYTSLIHS LIEESQNQQEKNEQ

2C

Linker oligopeptide: SEQ ID NO 2

S G G R G G S

2D

Region 604-615: SEQ ID NO 15 : S G K L I C T T A V P W (-12 residues)

Region 598-622: SEQ ID NO 16: L G I W G C S G K L I C T T A V P W N A S W S N K (-25 residues)

FIGURE 2

SEQ ID NO 13 and 14 (GP41 matrix) :

(V) Q540 A R Q L L S G I V Q Q Q N N L L L 556
(GTA) CAG GCC AGA CAA TTA TTG TCT GGT ATA GTG CAG CAG AAC AAT TTG CTG
R A I E A Q Q Q H L L Q L T V W G I K 574
AGG GCT ATT GAG GCG CAA CAG CAT CTG TTG CAA CTC ACA GTC TGG GGC ATC AAG
Q L Q A R I L A V E R Y L K D Q Q L 592
CAG CTC CAG GCA AGA ATC CTG GCT GTG GAA AGA TAC CTA AAG GAT CAA CAG CTC
L G I W G S S G K L I S T A V P W 610
CTG GGG ATT TGG GGT AGC TCT GGA AAA CTC ATT AGC ACC ACT GCT GTG CCT TGG
N A S W S N K S L E Q I W N H T T W 628
AAT GCT AGT TGG AGT AAT AAA TCT CTG GAA CAG ATT TGG AAT CAC ACG ACC TGG
M E W D R E I N N Y T S L I H S L I 646
ATG GAG TGG GAC AGA GAA ATT AAC AAT TAC ACA AGC TTA ATA CAC TCC TTA ATT
E E S Q N Q Q E K N E Q E L L E L D 664
GAA GAA TCG CAA AAC CAG CAA AAG AAT GAA CAA GAA TTA TTG GAA TTA GAT
K W A S L W N W F N I 675
AAA TGG GCA AGT TTG TGG AAT TGG TTT AAC ATA

FIGURE 3

SEQ ID NO 17

M	Q	A	R	Q	L	L	S	G	I	V	Q	Q	Q	N	N	L	L	18
R	A	I	E	A	Q	Q	H	L	L	Q	L	T	V	W	G	I	K	36
Q	L	Q	A	R	I	L	A	V	E	R	Y	L	K	D	Q	Q	L	54
S	G	G	R	G	G	S	S	L	E	Q	I	W	N	H	T	T	W	72
M	E	W	D	R	E	I	N	N	Y	T	S	L	I	H	S	L	I	90
E	E	S	Q	N	Q	Q	E	K	N	E	Q	E	L	L	E	L	D	108
K	W	A	S	L	W	N	W	F	N	I	T	N	W	L	D	H	H	126
H	H	H	H															

FIGURE 4A

SEQ ID NO 18

M	Q	A	R	Q	L	L	S	G	I	V	Q	Q	Q	N	N	L	L	18
R	A	I	E	A	Q	Q	H	L	L	Q	L	T	V	W	G	I	K	36
Q	L	Q	A	R	I	L	A	V	E	R	Y	L	K	D	Q	Q	L	54
S	G	G	R	G	G	S	S	L	E	Q	I	W	N	H	T	T	W	72
M	E	W	D	R	E	I	N	N	Y	T	S	L	I	H	S	L	I	90
E	E	S	Q	N	Q	Q	E	K	N	E	Q	E	L	L	E	L	D	108
K	W	A	S	L	W	N	W	F	N	I	T	N	D	H	H	H	H	126
H	H																	

FIGURE 4B

SEQ ID NO 19

M	Q	A	R	Q	L	L	S	G	I	V	Q	Q	Q	N	N	L	L	18
R	A	I	E	A	Q	Q	H	L	L	Q	L	T	V	W	G	I	K	36
Q	L	Q	A	R	I	L	A	V	E	R	Y	L	K	D	Q	Q	L	54
L	G	I	W	G	S	S	G	G	R	G	G	S	S	L	E	Q	I	72
W	N	H	T	T	W	M	E	W	D	R	E	I	N	N	Y	T	S	90
L	I	H	S	L	I	E	E	S	Q	N	Q	Q	E	K	N	E	Q	108
E	L	L	E	L	D	K	W	A	S	L	W	N	W	F	N	I	T	126
N	W	L	D	H	H	H	H	H	H									136

FIGURE 5A

SEQ ID NO 20

M	Q	A	R	Q	L	L	S	G	I	V	Q	Q	Q	N	N	L	L	18
R	A	I	E	A	Q	Q	H	L	L	Q	L	T	V	W	G	I	K	36
Q	L	Q	A	R	I	L	A	V	E	R	Y	L	K	D	Q	Q	L	54
L	G	I	W	G	S	S	G	G	R	G	G	S	S	L	E	Q	I	72
W	N	H	T	T	W	M	E	W	D	R	E	I	N	N	Y	T	S	90
L	I	H	S	L	I	E	E	S	Q	N	Q	Q	E	K	N	E	Q	108
E	L	L	E	L	D	K	W	A	S	L	W	N	W	F	N	I	T	126
N	D	H	H	H	H	H	H											134

FIGURE 5B

SEQ ID NO 21

														M	L	L	3	
R	A	I	E	A	Q	Q	H	L	L	Q	L	T	V	W	G	I	K	21
Q	L	Q	A	R	I	L	A	V	E	R	Y	L	K	D	Q	Q	L	39
S	G	G	R	G	G	S	S	L	E	Q	I	W	N	H	T	T	W	57
M	E	W	D	R	E	I	N	N	Y	T	S	L	I	H	S	L	I	75
E	E	S	Q	N	Q	Q	E	K	N	E	Q	E	L	L	E	L	D	93
K	W	A	S	L	W	N	W	F	N	I	T	N	W	L	D	H	H	111
H	H	H	H															115

FIGURE 6